center for scientific review

peer review notes

September 2009

In This Issue

You Met The Challenge —Thank You!

Many of you and your colleagues showed the tremendous vitality of the U.S. research community by submitting more than 25,000 grant applications for funding through the American Recovery and Reinvestment Act (ARRA). It was unprecedented, and reviewing all these applications was one of the most challenging tasks NIH and CSR have ever faced.

The most wonderful part of this story was how so many of you rallied when we asked for your help in reviewing these applications. A total of 20,000 of you said yes and reviewed your assigned applications in less than a month! Your generosity, commitment and extraordinary efforts provided the necessary expertise to review these applications so NIH could find and fund the best to advance the economy, science and health.

The Congress, President Obama and the American public put a lot of hope and trust in NIH and the scientific community, and you came through. All of us at CSR and NIH thank you.

Peer Review Enhanced

NIH successfully implemented the most sweeping enhancements to peer review in its 60-year history last round.

"Things went amazingly well," said CSR Director Toni Scarpa. "Our chairs and reviewers did an incredible job implementing these changes. It would have been impossible without the many inside and outside NIH who helped make it happen."

The launch of these peer review changes, coupled with the added task of reviewing so many ARRA applications, posed an exceptional challenge to NIH staff—but one met with equal parts of commitment and success. "We also owe a special thanks to Scientific Review Officers, who worked to keep review meetings consistent with enhancements to achieve this major milestone," Dr. Scarpa said.

The new scoring system worked well. Reviewers used the full range, and their scores were well-distributed with less "compression" than in the old scoring system. The new system is based on 1-9 scoring and assigns each criterion a value, along with an overall impact/priority score.

Other Key Changes

- Structured critiques are now aligned to the review criteria.
- New emphasis is placed on impact.
- Preliminary impact/priority scores are now used by CSR review groups to order their reviews and facilitate score calibration and decisions on which applications will not be discussed.

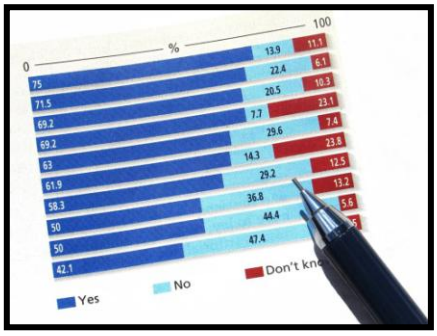
Points to Keep in Mind for New Reviewers

- Focus on critiquing the applications instead of going overboard mentoring the applicants.
- Write succinct, but specific, bullets when reporting on the strength and weakness of an application.
- Refrain from repeating information in the application in your critiques.
- Know the difference between "significance" and "overall impact." Significance addresses whether (and to whom) the project or work proposed is important; impact addresses the likelihood for this project to have an effect on the field, which means not only the importance of the goals but also the likelihood of success in achieving those goals. The overall impact/priority score is not, however, a summation of criterion scores.

Shorter applications better aligned to the review criteria: Starting with applications due January 25, 2010, and beyond, applicants will have new forms and instructions to apply for all NIH grants. The R01 grant application and most others will be shorter and better aligned to the review criteria. Detailed information has been posted in the *NIH Guide*: **http://grants1.nih.gov/grants/guide/notice-files/NOT-OD-09-149.html**.

Stay tuned for more details about enhancements soon to be publish in the *NIH Guide* and the Enhancing Peer Review Web site: **http://enhancing-peer-review.nih.gov/timelines.html**. NIH intends to assess and refine these changes by conducting a systematic study of reviewers, applicants, Scientific Review Officers (SROs) and program staff. In the meantime, please send any comments and suggestions to the following address: **EnhancingPeerReview@mail.nih.gov**.

CSR Studies 1-9 Scoring and New Review Practices



CSR recently assessed stakeholder satisfaction with using 1-9 scoring, reviewing applications online, using two–stage editorial reviews, and giving applicants a prebuttal option prior to final scoring of their application. CSR queried reviewers and staff who participated in pilot studies of these practices to enhance peer review. Their input will help CSR evaluate how well the system modifications can achieve their respective goals. Detailed reports are going online.

1-9 Scoring

NIH used 1-9 scoring for applications received for funding in fiscal year 2010 and also for applications received in response to American Recovery and Reinvention Act initiatives. The goal was to create a rating scale that realistically captures the range of quality in applications

that study sections review and to encourage routine use of the entire rating scale. Prior to its introduction, CSR polled reviewers in two study sections that piloted 1-9 scoring.

One-hundred percent of the respondents were satisfied with the new scoring system and the way it helped to identify scientific merit and impact. Respondents also said applicants may find the new scoring system useful for interpreting reviewer feedback. In addition, they said that other reviewers would not find it an additional burden, compared to the 1-5 system.

Asynchronous Electronic Discussion Meetings

Since 2006, CSR has used online threaded message boards to engage the best possible reviewers, especially when they do not wish to travel to meetings or when this online format represents the best way to review a group of applications. CSR recently polled reviewers and SROs to assess these Asynchronous Electronic Discussion (AED) reviews and make improvements.

Key Results: Nearly 60% of reviewers and 70% of the SROs said AED meetings were less of a burden than face-to-face meetings. All groups preferred AED to be used for smaller groups of applications per meeting—1-10 applications was most preferred; 11-20 applications was second. Thirty-six percent of reviewers would choose an AED meeting over a face-to-face meeting for their next meeting and a sizable number of reviewers (21%) expressed no preference.

Improvements: The level of reviewer participation was a concern noted by participants, who offered many suggestions for making improvements.

Two-Stage Editorial Board Reviews

NIH used two-stage, or editorial board, reviews to assess applications for several key grant mechanisms. In this type of review, two to three experts assess each application and submit their critiques online. Then, a second panel of reviewers (editorial board) with broad expertise examines the first critiques and applications, focusing on the impact of the proposed research and assigning final overall impact/priority scores.

- A majority of reviewers studied said they would participate in two-stage review in the future and that they would choose this review method for their own applications.
- The second stage reviewers said that they were well informed by the first-stage reviewers.

CSR plans to study the effectiveness of two-stage review for other grant mechanisms, such as Challenge grants.

Prebuttal Pilot

Within the two-stage review platform, CSR also assessed the value of giving applicants the opportunity to provide a prebuttal, so they could address factual errors in reviewer critiques prior to the second level of review. The study of prebuttal, which is just in the assessment stage, was a second pilot nested in a two-stage review study of small business applications.

Only a small percentage of prebuttals identified factual scientific errors, although 80 percent of applicants submitted prebuttals. Reviewers and SROs studied reported prebuttals rarely made a difference in reviews and that applicants often went beyond correcting just factual errors. CSR and NIH will revisit this issue after the current enhancements to peer review are

implemented and working well.

CSR posts reports on these and other CSR studies as soon as they are available on our new Evaluation Reports Web site: **http://www.csr.nih.gov/EvaluationReports/**.

NIH is working to assess trans-NIH peer review enhancements among internal and external constituencies, to include applicants, reviewers, SROs and program officers.

Staff and Reviewers Rally in Challenging Review Effort

It was one very late night in May. CSR and its reviewers were in the heat of their reviews of NIH Challenge grant applications for federal stimulus funds. It was precisely 2:40 a.m. when CSR's IRG Chief Dr. Noni Byrnes had a question and sent her query to five people. Within five minutes, all five responded: three SROs, another Integrated Review Group (IRG) chief and a support staff member.

Byrnes led the massive effort to review—in three months—nearly 20,000 Challenge applications for funding through the American Recovery and Reinvestment Act. "She and our team did heroic jobs," said CSR Director Dr. Toni Scarpa. "We had 90 days from the time of receipt to the time of releasing the summary statements. That's unprecedented, and the fact that it was all done efficiently and on time shows the incredible commitment of hundreds of staff members from across CSR and NIH as well as 15,000 reviewers."

Byrnes expanded on this point, saying that "one person could not possibly think of the many, many things that could have gone wrong—but that didn't, because there were so many heads together brainstorming and responding quickly to the various issues arising in a review effort this massive."

It helped that officials from CSR and NIH had developed plans ahead of time to be prepared for dealing with different submission levels from a few thousand to the nearly 20,000 that actually came. "We were ready," Dr. Scarpa said. "When we realized, after getting feedback from the community, that there would be so many applications, we used a system that was developed, tested and analyzed previously." This system helped NIH efficiently identify truly innovative science within a compressed review schedule. We specifically used two-stage editorial board reviews and enhancements developed through the enhancing peer review initiative: shorter applications, the structured critique template, and CSR's practice of reviewing applications in order of their preliminary overall impact/priority scores.

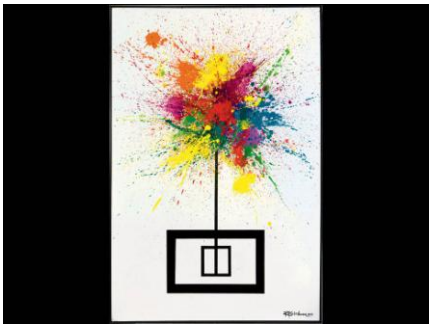
Of course, there were times when many had doubts. "I wasn't sure in the beginning that the stage-two reviewers would get enough information from the stage-one [mail] reviewers," said SRO Dr. Bonnie Burgess-Beusse, "but they did and the discussions went really well, with reviewers raising excellent points," she said. Dr. Lila Gierasch from the University of Massachusetts at Amherst agreed. "I had doubts many steps of the way, but in the end, having sat on the panel, I was quite impressed. The two levels of review worked quite well, and there was a remarkable performance at CSR to orchestrate the review of that many applications. And we found some truly exciting projects among the proposals we reviewed."

It was an incredible experience for the thousands of Challenge reviewers and CSR staff, as well as thousands of other reviewers and NIH staff who marshaled to review and process

applications for other grants funded through the American Recovery and Reinvestment Act. “We worked together,” said IRG Chief Dr. Mushtaq Khan, “lost a lot of sleep, working lots of hours, but we did it. We met the deadline and breathed a huge sigh of relief.”

Program and grants management staff members at the NIH Institutes and Centers have a lot of work to do now. They will join CSR staff and reviewers in looking forward to seeing the stimulus funds support research that will advance the economy, science and health.

Praise and Insight for the Transformative Research Program



The first round of reviews for transformative research grants is complete, and feedback from reviewers and applicants offers insights for improving the program and stresses the importance of encouraging innovative proposals in the future.

NIH created the initiative, known as Transformative Research Projects Program (T-R01), to respond to complaints from the scientific community that NIH was not identifying enough innovative science that would transform research fields.

NIH utilized a three-stage review process to better identify innovation. It recruited pioneers from various fields and other experts versed in the technical aspects of research. About a dozen eminent scientists, or “editors,” selected for their broad grasps of science identified the grant applications holding the greatest potential for transforming a given field. In the second stage, subject-matter experts submitted input on the applications. The editors then reviewed this input and the applications and met to make the final assessments.

“The editors tolerated a much higher level of risk, and they were not defending their own critiques, demonstrating that they were not heavily invested into these points of view,” said CSR Director Toni Scarpa. “They were more willing to give and take in the discussions.”

A Web-based assessment showed support among reviewers and applicants for the program, with some recommendations for how to improve it. This study queried the 720 applicants and eleven reviewers, with 62 percent of applicants and 63 percent of reviewers responding. The studies were fielded in February to May of 2009, before the applicants were notified on the status of their applications, but after reviewers in the first and third stages made final decisions.

The three-stage review process was specifically designed for the task, said Dr. Keith Yamamoto, from UCSF, who co-chaired one of the editorial boards. “The editors were selected for their ability to identify transformative ideas even outside their own primary knowledge base. They can ‘insulate’ the applicants from the content experts, some of whom may have created a prevailing paradigm and therefore might defend it reflexively . . . . Finding the right balance of expert reviewer and generalist editor input is critical to the success of the T-R01 mechanism.”

Some Study Highlights

- All third stage reviewers asserted that the subject experts “always” or “often” had the appropriate expertise to evaluate the technical aspects of the application. Their input was “very helpful” in arriving at a final score and it “sometimes” or “often” dramatically changed the initial stage one assessment of the proposals.
- Both applicants and reviewers opposed highlighting areas of NIH interest in the solicitation for grant applications, with reviewers saying that doing so increased the number of non-competitive applications and reviewer workload.

- Reviewers said that more than 50 percent of the applicants understood the requirements of the funding announcement, but fewer than 25 percent of applications proposed truly transformative research.

T-R01s sparked creativity and innovation: "We got what we wanted out of this effort [and] we have tweaks we'd like to make for the future," said Dr. Elizabeth Wilder, deputy director of NIH's Office of Strategic Coordination, which coordinated the T-R01 initiative. The T-R01 initiative itself, however, holds big promise for science innovation, she said. "There is a core need for dedicated funding to encourage science to go beyond the bounds of R01s."

"We believe scientists will come to love this program and that we will get more applications that are outside of the box. This round there were 700 applications, but we expect much more in the future." Congress demonstrated its support of this type of high-risk, high-reward science by continuing its fiscal year 2009 funding into fiscal 2010. The next T-R01 deadline has not officially been set, but Dr. Wilder expects it will fall around the same time as the first deadline, which was in January.

Finding Homes for 'Orphan' Applications

A new review approach is needed to ensure that "orphan" applications get fair assessments in a review environment that maximizes competition and encourages science. This was the conclusion of two CSR studies.

We define "orphan" applications as those with ideas that do not easily fit into one of CSR's several hundred study sections or they represent less than 5 percent of the applications in a study section that are assigned to an NIH institute or center (IC). CSR conducted two studies in response to a request from the Peer Review Advisory Committee (PRAC) to examine how well CSR treats these types of applications.

IC orphans represent a small percentage of overall applications, according to the first study, which reviewed R01 grant applications submitted in 2007. That year, there were a total of 2,200 orphans, representing 8 percent of the total or 700 applications a cycle. SROs and program officers were queried in a second study. They identified fewer orphans, about 2 percent, though no orphan application was identified by both an SRO and a program officer.

Outcome Differences: While all ICs have orphans, large ICs tend to have more. Review outcomes varied for orphans in small ICs and those in larger ones, and for new vs. established investigators. But Dr. Donald Schneider, CSR division director, told PRAC that outcome differences are not predictable and they often are "not dramatic." He also said that the data were "only suggestive, given the small sample size."

PRAC members discussed employing diverse expertise and using cross assignments to review these applications to ensure they are reviewed appropriately. For more information on this presentation, visit the PRAC Web site: **http://grants1.nih.gov/grants/peer/prac/index.htm**.

Subscribe to Peer Review Notes:
http://cms.csr.nih.gov/NewsandReports/PeerReviewNotes
Send comments or questions: PRN@csr.nih.gov.

Center for Scientific Review
National Institutes of Health
U.S. Department of Health and Human Services